

学校编码: 10384 分类号密级

学号: 30620141153070 UDC

厦门大学

硕士学位论文

# 个性化推荐系统评价指标体系的实证研究

## Empirical Research of Dimensions and Metrics for Evaluating Personalized Recommender System

曾洁

指导教师姓名: 曾秀芹副教授

专业名称: 新闻与传播

论文提交日期: 2016 年 4 月

论文答辩时间: 2016 年 5 月

学位授予日期: 2016 年 6 月

答辩委员会主席: 王霏

评阅人: 柯学、高丽华

2016 年 6 月

## 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（ ）课题（组）的研究成果，获得（ ）课题（组）经费或实验室的资助，在（ ）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

曾 浩

2016 年 5 月 25 日

## 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（        ） 1. 经厦门大学保密委员会审查核定的保密学位论文，  
于        年        月        日解密，解密后适用上述授权。

（        ） 2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

曾 浩

2016 年    5 月    25 日

## 摘要

随着互联网竞争日益激化，各大网站在个性化服务方面不断推新除陈，通过优质的推荐服务来争取用户。近年来，以淘宝网、京东网在内的国内各大网络巨头不断优化自身的推荐系统服务，除业界重视外，推荐系统也受到众多学科和研究者的关注。以往有关推荐系统的研究呈现出技术导向的特征，以消费者为导向的研究较为匮乏，然而辅助消费者的决策是推荐系统的主要目标，虽然关于推荐系统的评价指标体系的研究尚有一些探讨，但评价体系还不够完整，对符合个性化推荐系统特点的评价指标体系定量评价方法研究仍是一个开放性的问题。

本文立足消费者视角，通过问卷调查方法，从实证角度研究个性化推荐系统的评价指标体系，在充分地讨论相关文献的基础之上，本文总结了个性化推荐系统的评价指标，即准确度、熟悉度、偏好表现、感知易用性等，并从这些评价指标出发，构建了一套具备扎实理论基础的个性化推荐系统的评价指标体系。这次的主要创新点是将结构方程模型加入到了个性化推荐系统评价体系中，我们得出了如下结论：个性化推荐系统的评价指标体系有五大因子，它们分别是：“内容评价”、“互动评价”、“技术评价”、“系统评价”、“主观态度评价”。

关键词：个性化推荐系统;评价体系;实证研究

## Abstract

With the increasing fierce of each domestic great website, more and more Internet websites have pursued the excellent recommendation service. In recent years, each great domestic network tycoon including Taobao.com, Jingdong.com optimizes its recommendation system service constantly. Except for the emphasis in the circle, recommendation system also receives the attention of numerous disciplines and researchers. Previous research on recommendation system has shown the technology oriented characteristics, consumer-oriented research is more deficient, but the decision that assists consumers is the main goal of the recommendation system. Although some discussions exist on the research of the evaluation index system of the recommendation system, evaluation system is not complete, research on the quantitative evaluation method with the characteristics of personalized recommendation system is still an open question.

This paper is based on the perspectives of consumers, through the method of questionnaire investigation, it studies the evaluation index system from the empirical angle. On the basis of fully discussing related literature, this paper summarizes the evaluation index of personalized recommendation system, or accuracy, familiarity, preference performance, perceptive usability, and it starts from these evaluation indexes to construct a set of evaluation index system of the personalized recommendation system with the solid theoretical basis. This paper applies structural equation model into the quantitative evaluation research of personalized recommendation system creatively. The following conclusion is obtained: there are five great factors in evaluation index system of personalized recommendation system, which are “content evaluation”, “interactive evaluation”, “technology evaluation”, “system evaluation”, “subjective attitude evaluation”.

**Key Words:** Personalized recommendation system; Evaluation system ;Structure equation; Empirical research

## 目录

摘要.....	I
第 1 章 绪论.....	1
1.1 研究背景.....	1
1.2 研究目的和意义.....	2
1.2.1 研究目的.....	2
1.2.2 研究意义.....	3
1.3 研究框架.....	3
第 2 章 文献综述.....	6
2.1 推荐系统的产生和发展.....	6
2.2 推荐系统的定义.....	7
2.3 推荐系统的类型.....	9
2.3.1 按推荐算法分类.....	9
2.3.2 按个性化程度的分类.....	10
2.3.3 按自动化程度的分类.....	10
2.3.4 按持久性程度的分类.....	11
2.4 推荐系统的推荐方式.....	11
2.5 推荐系统的评价.....	17
2.5.1 基于推荐算法的评价指标.....	17
2.5.2 基于系统角度的评价指标.....	19
2.5.3 基于用户角度的评价指标.....	20
第 3 章 研究设计.....	23
3.1 推荐系统评价指标体系研究的假设理论模型.....	23
3.2 推荐系统评价指标的界定.....	24
3.2.1 内容评价.....	24
3.2.2 互动评价.....	27
3.2.3 技术评价.....	28
3.2.4 系统评价.....	28

3.2.5 主观态度评价.....	28
<b>第 4 章 研究方法.....</b>	<b>33</b>
4.1 问卷调查.....	33
4.1.1 调查目的.....	33
4.1.2 问卷的制定.....	33
4.1.3 问卷的执行.....	34
4.1.4 问卷的收集方式.....	35
4.2 数据分析方法.....	35
<b>第 5 章 数据分析.....</b>	<b>36</b>
5.1 预试调研的结果分析.....	36
5.1.1 内部一致性检验.....	36
5.1.2 量表结构效度检验.....	36
5.2 正式研究的结果分析.....	39
5.2.1 内容评价测量模型.....	39
5.2.2 互动评价测量模型.....	41
5.2.3 技术评价测量模型.....	42
5.2.4 系统评价测量模型.....	43
5.2.5 主观态度评价测量模型.....	44
5.2.6 二阶因子分析.....	45
<b>第 6 章 结论与讨论.....</b>	<b>49</b>
6.1 个性化推荐系统的评价指标体系结论与讨论.....	49
6.2 对推荐系统和商家营销实践的启示.....	50
6.3 研究局限与未来研究建议.....	51
<b>参考文献.....</b>	<b>53</b>
<b>附录一：个性化推荐系统评价指标体系预调研问卷表.....</b>	<b>62</b>
<b>附录二：个性化推荐系统评价指标体系正式调研问卷表.....</b>	<b>66</b>

# Contents

<b>Abstract.....</b>	<b>II</b>
<b>Chapter1 Introduction.....</b>	<b>1</b>
<b>1.1 Research Background.....</b>	<b>1</b>
<b>1.2 Purpose And Significance Of Research.....</b>	<b>2</b>
1.2.1 Purpose Of Research.....	2
1.2.2 Significance Of Research.....	3
<b>1.3 Framework Of Research.....</b>	<b>4</b>
<b>Chapter2 Literature Review.....</b>	<b>6</b>
<b>2.1 Formation And Development Of Recommender Sysytem.....</b>	<b>6</b>
<b>2.2 Definition Of Recommender Sysytem.....</b>	<b>7</b>
<b>2.3 Type Of Recommender Sysytem.....</b>	<b>9</b>
2.3.1 The Recommended Classification Algorithm.....	9
2.3.2According To The Classification Of Individuation Degree.....	10
2.3.3According To The Classification Of Individuation Degree.....	11
2.3.4According To The Classification Of The Degree Of Persistence.....	11
2.4 Strategies Of Recommender Sysytem.....	12
2.5 Evaluation Of Recommender Sysytem.....	18
2.5.1 Based On The Evaluation Index Of Recommendation Algorithm.....	18
2.5.2 Based On The Evaluation Index System Point Of View.....	20
2.5.3 Based On The User Perspective Evaluation Index.....	21
<b>Chapter3 Research Design.....</b>	<b>24</b>
<b>3.1 Hypothesis Theoretical Model.....</b>	<b>24</b>
<b>3.2 Definition Of Evaluation Metrics.....</b>	<b>26</b>
3.2.1Content Evaluation.....	26
3.2.2Interactive Evaluation.....	28
3.2.3Technical Evaluation.....	29
3.2.4 System Evaluation.....	29
3.2.5Subjective Attitude Evaluation.....	31
<b>Chapter4 Research Methods.....</b>	<b>34</b>



<b>4.1 Questionnaire Survey.....</b>	<b>34</b>
4.1.1 Research Purpose.....	34
4.1.2 The Establishment Of the Questionnaire.....	34
4.1.3 The Implementation Of The Questionnaire.....	35
4.1.4 Methods The Questionnaire.....	36
<b>4.2 Data Analysis Methods.....</b>	<b>36</b>
<b>Chapter5 Data Analysis.....</b>	<b>38</b>
<b>5.1 Pre Research Results.....</b>	<b>38</b>
5.1.1 Internal Consistency Check.....	38
5.1.2 Scale Structure Validity Test.....	38
<b>5.2 Formal Research Results.....</b>	<b>41</b>
5.2.1 Content Evaluation Measurement Model.....	41
5.2. 2 Interactive Evaluation Measurement Model.....	43
5.2.3 Technical Evaluation Measurement Model.....	44
5.2.4 System Evaluation Measurement Model.....	45
5.2. 5 Subjective Attitude Evaluation Measurement Model.....	46
5.2.6 A Second Order Factor Analysis.....	47
<b>Chapter6 Discussion.....</b>	<b>50</b>
<b>6.1 Personalized Recommendation System Of Evaluation Index System Of The Conclusion And Discussion.....</b>	<b>50</b>
<b>6.2 Contribution And Implication.....</b>	<b>51</b>
<b>6.3 Research Limitations And Prospects.....</b>	<b>52</b>
<b>Reference.....</b>	<b>54</b>
<b>Appendix 1.....</b>	<b>63</b>
<b>Appendix 2.....</b>	<b>67</b>

## 第 1 章绪论

### 1.1 研究背景

互联网出现后，随着网上内容的增加，网络带来庞大的信息流，可互联网不像图书馆搞个书目索引就行，于是出现了搜索引擎帮助用户在茫茫互联网上找到他们感兴趣的东西，但条件是你必须知道你想要什么，然后提取成关键字去搜，所谓信息检索。

随着信息领域相关技术的飞速发展，互联网逐渐充斥在社会的每一个角落，改变了人们一直以来的生活方式，人类进入了一个全新的高度信息化的时代。尤其是进入 Web 2.0 时期以来，随着社会化网络媒体的异军突起，互联网用户既是网络信息的接受者，也是网络内容的生产者，互联网络中信息的数量是非常庞大的，而用户对于互联网络信息的辨识能力是有限的，使得在互联网中找寻有用信息的成本巨大，产生了所谓的“信息过载”问题。这个时候搜索引擎的短板出现了，用户可能面对的情况是有时自己都不知道如何用关键字去描述自己想要的信息或者东西，亦或是知道关键字，然而搜索引擎动辄返回几十万个结果，面对如此庞大的选择集用户依然无所适从，对于信息冗杂而无法进行获取有效信息这个问题，个性化推荐应运而生。

亚马逊（Amazon）的 CEO Jeff Bezos 曾经说过，如果有一百万的用户亚马逊将为这一百万的用户提供一百万个网站。在亚马逊诞生至今，网站立志于这项技术的开发，并且整个公司在很大程度上基于非常大的支持。现如今亚马逊的这项技术已经非常的成熟了，能够为广大的用户定制个性化的推荐，从效果看已经被广大用户信赖并且使用。

随着电子商务规模的不断扩大，商品的品类和个数呈递增态势，越来越多的电子商务网站意识到了推荐系统在辅助用户决策上的重要地位，比较著名的应用推荐系统的电子商务网站有亚马逊、ebay、淘宝网、天猫商城等。在这些电子商务平台中，网站提供的商品数量不计其数，网站中的用户规模也非常巨大。据不完全统计，天猫商城中的产品数量已经超过了 4 000 万。在如此庞大的电商网站中，用户根据自己的购买意图输入关键字查询后，会得到很多相似的结果，用户在这些结果中也很难区分异同，也难于选择合适的物品。目前比较成功的电子商务网站中，每一个都在挖掘客户的数据，并且希望能够用算法为

用户提供相应的商品，以此来推动自身网站的发展。

虽然推荐系统已经被成功运用于很多大型系统及网站，就现在来看，在大数据之下的推荐系统，对于用户的产品推荐显然是不能够胜任的，这是由于推荐系统所面临的问题是十分复杂的，而进行单纯的算法计算并不能够让客户满意，而且对于商家来说也是怨声载道的。例如，用户数目越来越多产生的大数据本身具有的复杂性、不确定性和涌现性给推荐系统带来诸多新的挑战，基于实时模式的推荐在大数据集下面临着严峻考验，用户难以忍受超过秒级的推荐结果返回时间，在这样的情况下，推荐系统亟需建立一个从以往单纯以精确性为中心到综合考虑精确性、多样性和新颖性等因素的评估体系。正是基于这样的认识，笔者在梳理了相关的推荐系统的研究文献基础之上，展开了从用户角度切入对推荐系统评价的研究。

## 1.2 研究目的和意义

### 1.2.1 研究目的

这几年，关于推荐系统的推荐算法的计算在学术界已经成为一个热门的话题，这个问题大部分集中在纯学术领域，而对于算法之外的问题却鲜有研究，其实这种单纯的学术上的研究不能够从根本解决问题，还需要从用户的角度入手，尤其是从市场营销角度切入来研究用户如何感知推荐系统的方向还有待深入。用户对推荐技术接受的主观态度无论是对网络技术的研究者还是实践者是一个值得关注的重要问题。这是因为对于用户来说的好处包括更高的查找效率，更加优惠的产品，更多的购买决策的信心，以及一个潜在的发现新东西的机会。对于营销者来说，这项技术可以推荐给用户可能购买的产品，增加他们的整体满意度和忠诚度以及用户留存和推荐给朋友的可能性。因此，评价用户感知的推荐系统可以帮助开发人员和营销人员更准确地了解用户，这将反过来，有助于提高系统的准确预测一个特定的推荐。

基于此，本文提出以下两个研究目的：

(1) 推荐系统研究的新视野，从新的用户的角度来认识和厘清如何评价推荐系统这个问题。对推荐系统的评价问题是一个复杂的概念，以往的研究从不同角度的作了不同的解释来充实推荐系统评价方面的成果，也就是说，这也使得推荐系统的评价方面的研究较为混乱，对推荐系统评价的维度和指标的相关

研究虽然多，但没有得到衡量推荐系统评价的一张通用的评价量表，而这恰恰是互联网网站，尤其是电子商务平台迫切想要了解自身推荐系统在用户心中重要性的一个障碍。本研究将通过实证研究的方法来获取推荐评价指标体系的组成部分，并通过结构方程进行验证，经研究所获得的评价量表可以使平台管理者随时知道用户对推荐系统的看法，并对推荐系统进行不断的修正，以期获得更好的使用效用并带来更多的有形和无形资产。

(2) 以往的研究推荐系统的文章，较多涉及到机器学习、数据挖掘、信息检索等领域研究人员有关用户偏好的引出和预测算法的改进，该类研究在很长一段时间内在推荐系统相关研究中占主流，对于实证研究是比较忽视的。所以本研究不但采用了问卷调查这种方法，并且还有结构方程的数据分析方法，进而建构推荐系统评价指标的概念模型。

### 1.2.2 研究意义

第一，理论意义。本论文试图从用户的角度认知推荐系统，对推荐系统评价指标构成要素即推荐系统评价指标的概念模型或概念维度进行实证的探索性研究。本文通过对用户的问卷调查，运用统计方法探索推荐系统评价指标的构成维度和获得推荐系统评价指标测量量表。同时，反映出提高推荐系统认可度的具体途径，为电子商务推荐系统在除了算法开发之外如用户界面设计和可用性等方面提供指导，网站所有者们可以在推荐系统设计、开发过程中加以控制，促进用户对推荐系统的接受、使用。

第二，实践意义。在理论上构建推荐系统评价指标外，这对了解我国目前的电商平台推荐系统评价现状具有十分重要的实践意义，为我们了解具体的推荐系统评价打开了一扇窗口。优质的推荐系统对于电商平台来说至关重要，不仅能够给平台本身带来实实在在的经济效益，更能带来更多的无形资产。从这个意义上来说，电商平台了解推荐系统在用户心目中的状态就显得非常必要，本文希望从实际的角度，给予这个问题一个好的答案。

## 1.3 研究框架

本文在研究中主要有六章，其主要的思路如图 1 所示。

第 1 章绪论，阐述本文的选题背景，提出研究目的和意义并对研究框架进行梳理。

第 2 章文献综述，以第一章的框架为目标将其中的概念进行了阐述，对于整个文章中所要涉及的专业内容做出了非常详细的研究，并且提出了相应的假设。对于研究中所使用的变量的测算提出了定量计算的方法。在对假设进行阐述的过程中，打下了坚实的理论基础。

第 3 章研究设计，针对第二章进行了非常详细的研究，并且结合了数据进行研究，然后提出了相应的假设并对推荐系统评价指标体系的相关变量进行定义和量化。

第 4 章研究方法，对本文调研实施和研究方法作一个具体的、全面介绍。最后，实施前期的预调研，再通过信度检验和因子分析得出本次调研的正式问卷调查表。

第 5 章数据分析与假设检验，使用 SPSS22.0 用于数据描述性和探索性因子分析，使用 AMOS22.0 用于验证性因子分析和结构方程探讨。

第 6 章研究结论。针对前两章的内容进行分析并依据分析结果对改善顾客对推荐系统的评价的营销建议，对于本研究的局限和未来的研究展望也作出了相应的探讨。

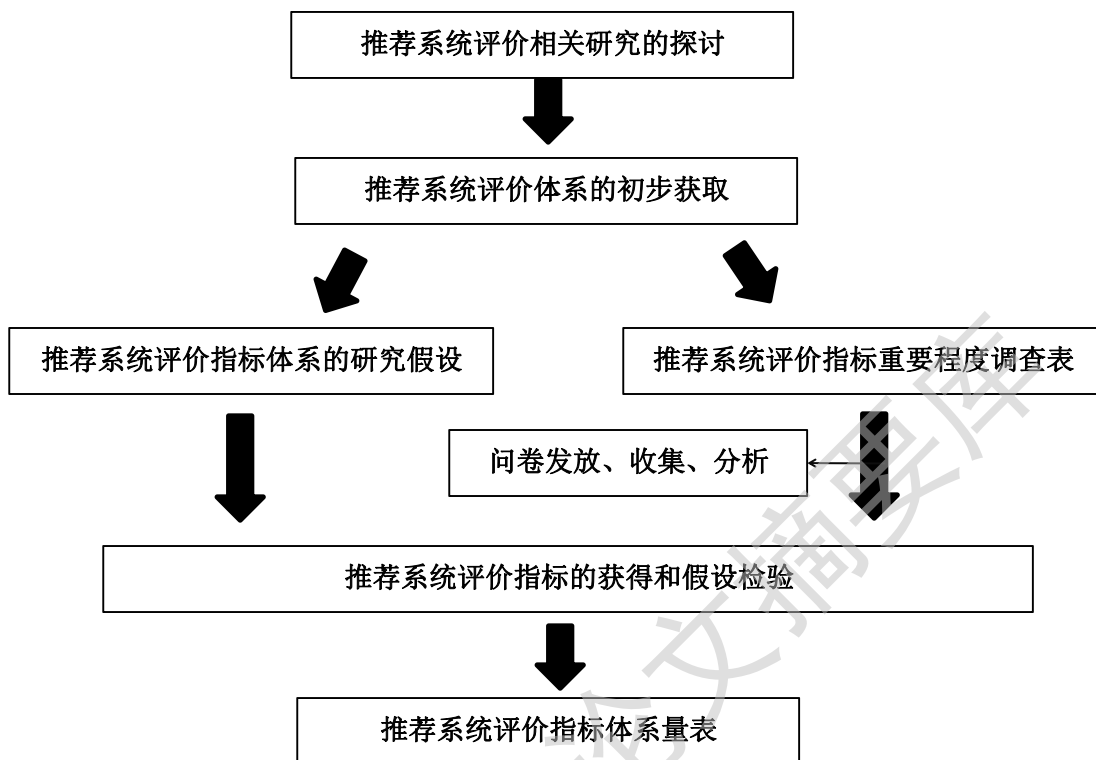


图 1 个性化推荐系统评价指标体系的实证研究的研究思路

## 第2章文献综述

### 2.1 推荐系统的产生和发展

在实际生活中，如果想要买零食，我们可以在便利店里慢慢的查找，由于店面有限，即便是将所有的商品看完也不会用到很多的时间。在比较大的超市中，用户可以借助于分类信息来找到自己的产品。但是，当面对更大选择的网络网站时，商品和商家的数量一下子变的巨大了。用户即便是有时间也不能够一一查看相应的商品。在这个时候，就需要借助于搜索引擎来对于自己所需要的东西进行查找。

当你需要面对很多选择的时候，正确的挑选就显得非常重要。当你有一百个同类型的商品需要进行挑选，这个时候你将变得非常的手足无措，而且由于挑选的难度很大，你会感觉非常困难。在互联网时代，这种状况变得更加糟糕，你将看到更多的商品，这些商品如果一个一个仔细看的话，你发现一个很不好的局面就是你根本无从选择。在巨大的选择集面前，人们通常是根本无从下手的，这不是选择恐惧症，而是一种处于本能的恐惧。网络在带给我们方便的时候，也将这种恐惧发送给我们。那么，究竟如何解决这样的一种困局呢？如何才能快速的找到想要的东西呢？

在这个信息时代，知识和信息的数量成倍的增加，人们每天有意无意能够接触到成千上万的信息，每一条讯息如果都看的话，那么人们将没有时间去做别的事情。在这个时代不是人主导信息的年代，信息已经犹如一个怪兽向人们袭来，人们无法阻挡信息的涌入，能够真正快速找到有用的信息已经成为了人们迫切的需要，很多时候信息太多，导致用户无法获得真正对自己有用的信息，对信息的有效使用率反而降低了，另外一方面真正有价值的信息也很难在垃圾信息中脱颖而出。

推荐系统就是一个根据用户的行为习惯测算出用户需要的信息或商品并进行推荐的一种方式。这个概念是 1995 年在美国人工智能协会（AAAI）上提出的，当时 CMU 大学的教授 Robert Armstrong 提出了这个概念，并推出了推荐系统的原型系统——Web Watcher。在同一个会议上，在美国名校斯坦福大学中，科研人员们做出了第一个基于自动化的计算机推荐系统 LIRA1。

推荐系统通过学习用户行为的模式，来探索某个人在从未体验过的物品集

合中将会喜欢的东西，极大地改变了人们对产品、信息以及其他用户的发现方式。21 世纪以来，推荐系统的研究与应用随着电子商务的快速发展而异军突起，各大电子商务网站都部署了推荐系统，其中 Amazon 网站的推荐系统比较著名。有报告称，Amazon 网站中 35% 的营业额来自于自身的推荐系统。随着推荐系统显示出了其巨大的商业价值和无限的发展潜力，它被业界寄予了很大的希望，希望能够解决信息太多给人们带来的困扰和麻烦，并且能够让人们可以更有效率的使用互联网。在我国国内很多互联网公司都开始了对于推荐系统的研究，这也研究已经成为电商领域的一各发展的重点，人们热衷于对于智能互联网推荐系统的研发。

当今中国互联网主要公司和其与推荐有关的产品和服务，包括：

- ① 淘宝、京东等电商网站的个性化产品推荐；
- ② 爱奇艺、土豆网等视频网站的视频推荐；
- ③ 酷狗、豆瓣音乐等音乐网站的音乐推荐；
- ④ 新浪微博、腾讯微博等社交网站的好友推荐；
- ⑤ 百度文库、鲜果网等网站的个性化阅读；
- ⑥ 各种个性化广告。

推荐系统诞生后，学术界对于互联网推荐系统的重视程度也是非常高的，在 1999 年美国就开始进行了针对基于电子商务的互联网推荐系统的国际会议，其会议内容体现了当前推荐系统研究的热点和发展趋势。最近的 10 年间，学术界对推荐系统越来越重视。目前为止，数据库、数据挖掘、人工智能、机器学习方面的重要国际会议都有大量与推荐系统相关的研究成果发表。

## 2.2 推荐系统的定义

在以往研究中，研究者对推荐系统没有形成统一的命名和定义，如交互式决策助手 (Castagnos, Jones, & Pu, 2009) (Häubl & Trifts, 2000a) (Murray & Häubl, 2008)、购物机器 (Hostler, Yoon, & Guimaraes, 2005; Pedersen, 2000)、产品比较代理 (Wan, 2005; Wan, Menon, & Ramaprasad, 2007)、推荐代理 (Xiao & Benbasat, 2007) 和推荐系统 (Cazella, Reategui, & Alvares, 2006; Schafer, Konstan, & Riedl, 1999) 等。

学者们针对推荐系统给出了很多的定义：



Degree papers are in the “[Xiamen University Electronic Theses and Dissertations Database](#)”.

Fulltexts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to [etd@xmu.edu.cn](mailto:etd@xmu.edu.cn) for delivery details.